

Inspect and analyze Waze data

Done by Charly Moreno (Data Analyst, 10/12/2023)

Overview

Investigate and understand the data provided regarding waze user's behavior. This data is a data extract that is planned to be used in a retention model to improve user retention.

Objective

The goal is to use a dataframe constructed within Python, perform a cursory inspection of the provided dataset, and inform team members of the findings.

Results

- The data shows 18% churned vs 82% retained users
- The labeled variable has 700 missing values.
- The maximums of these two variables are weird:
 - Driven_km_drives: 21,000 kms
 - Duration_minutes_drives: 263 hours per month
 - That's 700 kms per day, 87.5 kms per hour of drive (8 hours driven per day as well).
- Reviewing churned users, the driven_days per month variable is 6 vs 14 for retained. That implies:
 - Churned users with median_kms_driven_per_day of 608 kms (that's 8 hours at 76 km/h speed),
 - Retained users had a median of 247 kms.
 - Median drives_per_day also higher in the churned users group (8.3) vs the retained ones (3.3).

Next Steps

The analysis indicates that the type of users were serious drivers, so:

- It is suggested to ask for a more robust data set (with info from multiple months using random sampling).
- Add new variables for better profiling: Type of vehicle, age, sex and location at start / end of the session.